# Ecohydrological Process Networks and Phenology: Applying NEON and NPN Data Products

RISE and PROSE 2013
University of Arizona
Tucson, Arizona
11 October 2013

Macrosystems Biology Grant 2013-#1241960 (Conclusions are those of the authors and not necessarily the NSF)

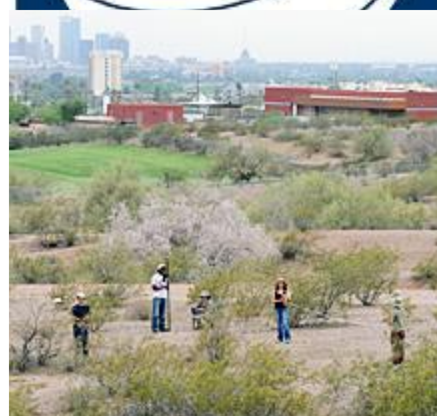
Benjamin L. Ruddell[1 2]
Dan Childers[2 3]

[1]Assistant Professor, College of Tech. and Innovation
[3]Professor, School of Sustainability
[2] Senior Sustainability Scientist, Global Inst. of Sustainability
Arizona State University

bruddell@asu.edu
480-727-5123

# The last 20 years have seen a Modeling and Observation Revolution for the Earth System

- Multi-scalar systemic observations & modeling
- Urban and socio-economic observatories like CAP-LTER
- How to design observatories?
- How to model systems?
- How to use all this data?
- By habit we work reductionistically, but this is complex systems science.





# Networks and Complex Adaptive Systems

- Complex Adaptive Systems (CAS's) involve multiple node types connected via multiple directional-weighted networks that create feedback and self-organizing control under dynamical conditions.
- Reductionism is inadequate for CAS's.
- In 1998 with 'Small World' networks (Watts & Strogatz, 1998) we realized that systemic network topology and behavior matter.
- But, standard network theory (and our thinking) focuses on social and communication networks, which are limited special cases.
- We need new network theories, especially for Process Networks.
- In this talk we will:
  - Briefly Introduce Process Network Concepts
  - Give examples of Dynamical Process Networks derived from Flux Tower and Phenology data
  - Introduce some hypotheses we are testing, and the NEON and NPN data products we are using to test them.

# How Reductionism Views a Process Network

A simple deterministic system is reduced so that independent variable $Y_1$ uniquely controls process X;

$$X = f\,(\,Y_1\,).$$

In this case we think we know what X and Y are and how to quantify them.

Experiments isolate the effect of Y on X and quantify *f* over a range of scales


This is a process network graph

- Nodes have types
- Connections have types
- Connections have direction and weight
- Connections may have rules or functions

Already, simple network theory does not apply

# How Reductionism Views a Process Network

But, there are other factors and subsystems, so process X is often a function of several Y's:

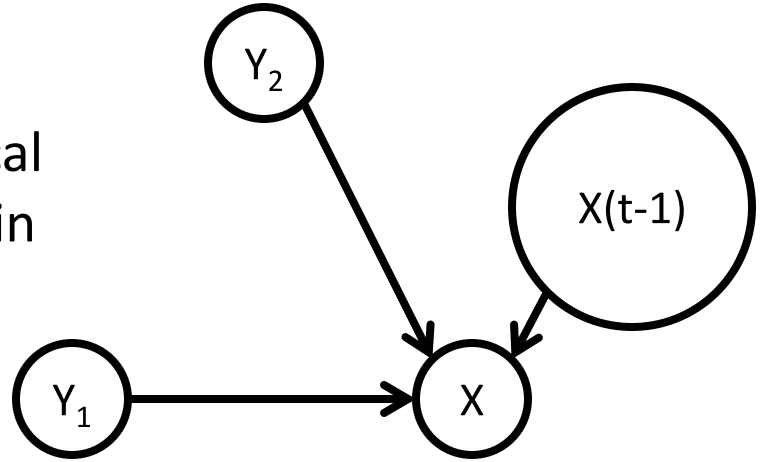
$X = f\,( Y_1, Y_2, \ldots Y_N\,).$

The typical conservation equation for a physical system includes the *history* or *position* of X in addition to *control* inputs:

$X(t) = X(t\text{-}1) + Y_1(t) + Y_2(t)$

Many systems are *approximately* like this, at least in the *net* sense and for a narrow range of *scales and assumptions*.

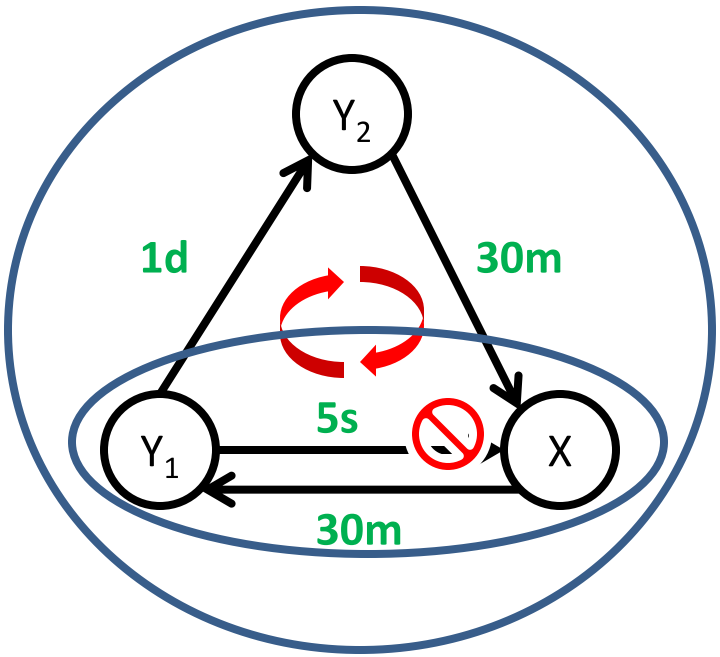
Mass and energy flow networks governed by conservation equations are good examples of this type of process network.



# How Complex Systems Science Views a Process Network

Complex systems generally feature coupling and *feedback* between many nodes, producing self-organizing subsystem behavior, and/or *thresholds* where key couplings turn on and off and qualitatively different system states emerge (Kumar 2007, Liu et al. 2007).

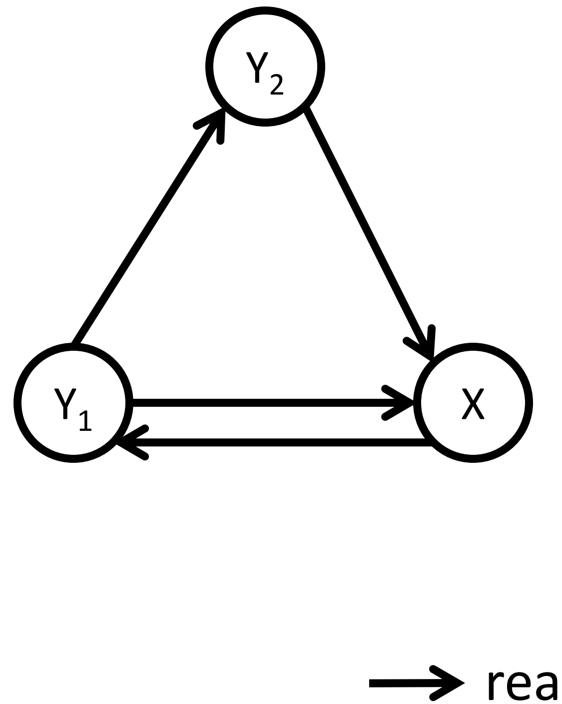
- Hierarchies of self-organizing subsystems can emerge via feedback.
- Connections have characteristic timescales at which processes operate.
- Connections have a *type, direction,* and *strength* (and possibly follow rules)
- In a *multitype network,* connections and nodes may be qualitatively different.




A ***Process Network (PN)*** is a network of feedback loops and the associated timescales that depicts the magnitude and direction of flow between the different subsystems. The PN graph itself defines the system state. (Ruddell and Kumar, 2009a)

# Flavors of Process Networks

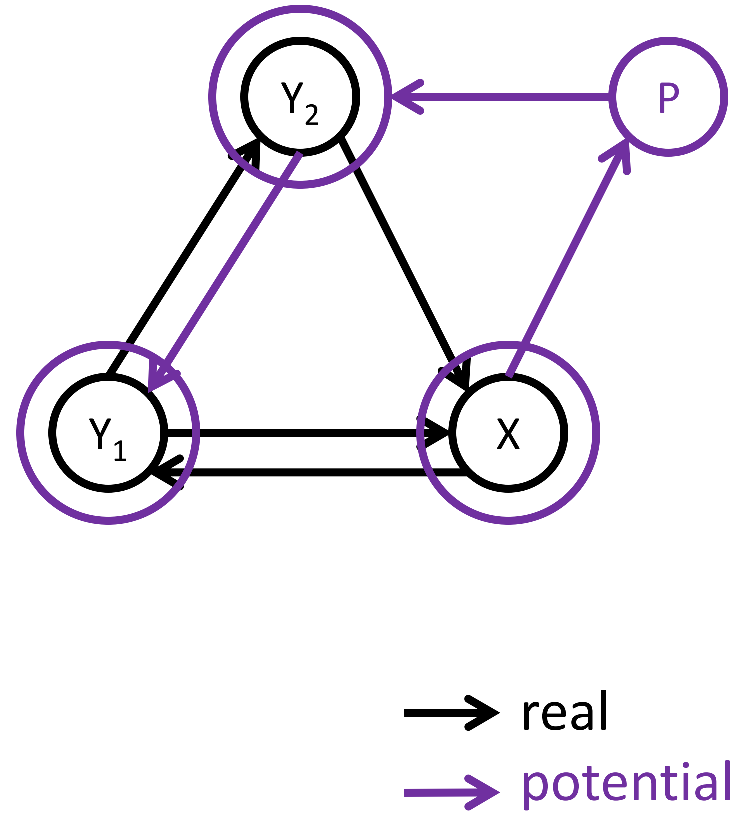
The *real* process network is what we usually think about. But we rarely (never?) know what that is with precision, especially at scales we cannot observe.



# Flavors of Process Networks

The *real* process network is what we usually think about. But we rarely (never?) know what that is with precision, especially at scales we cannot observe.

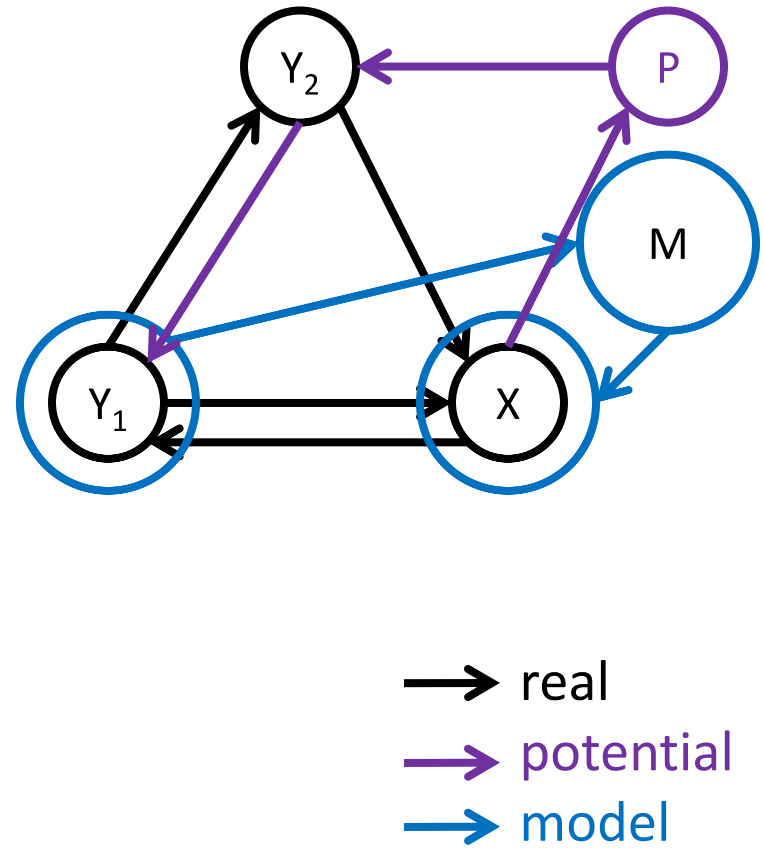
There is also the *potential* process network which includes all connections and nodes that could possibly ever exist, including those that do currently exist. This is even more difficult to establish.



# Flavors of Process Networks

The *model* process network is an approximation for modeling or observational purposes

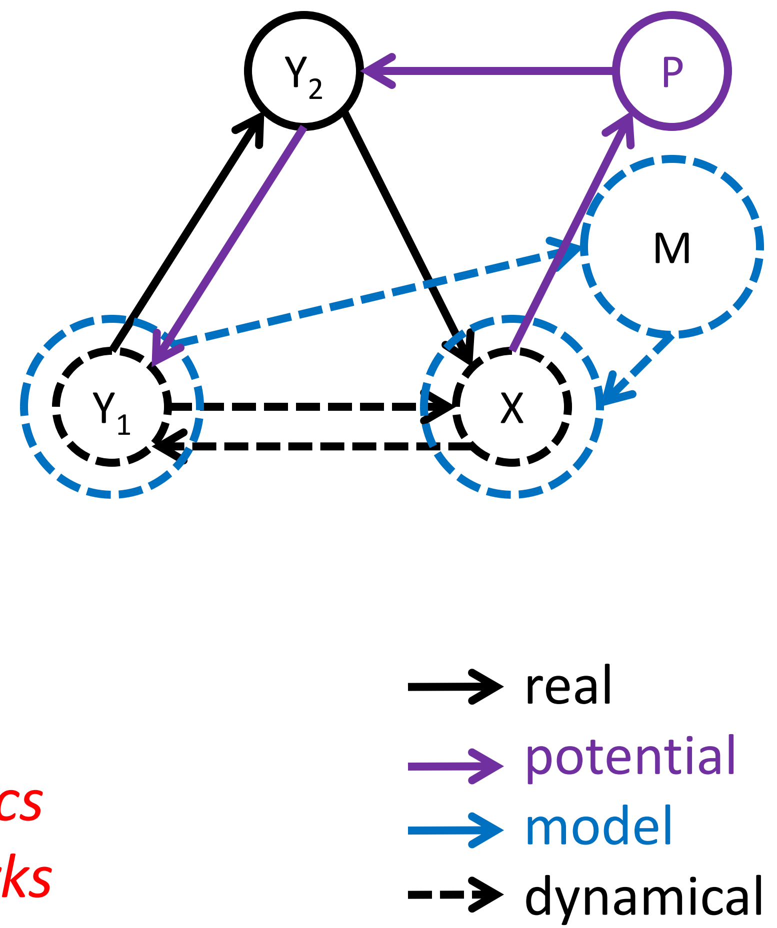
- uses measurable nodes and connections
- uses aggregated or simulated nodes *M*
- focuses on node(s) of interest *X* at a specific space-time scale
- aggregates real nodes that share a space and type to reduce detail, where possible
- Usually distorts the true system structure and behavior, but hopefully not too much
- Used to predict or to test hypotheses



# Flavors of Process Networks

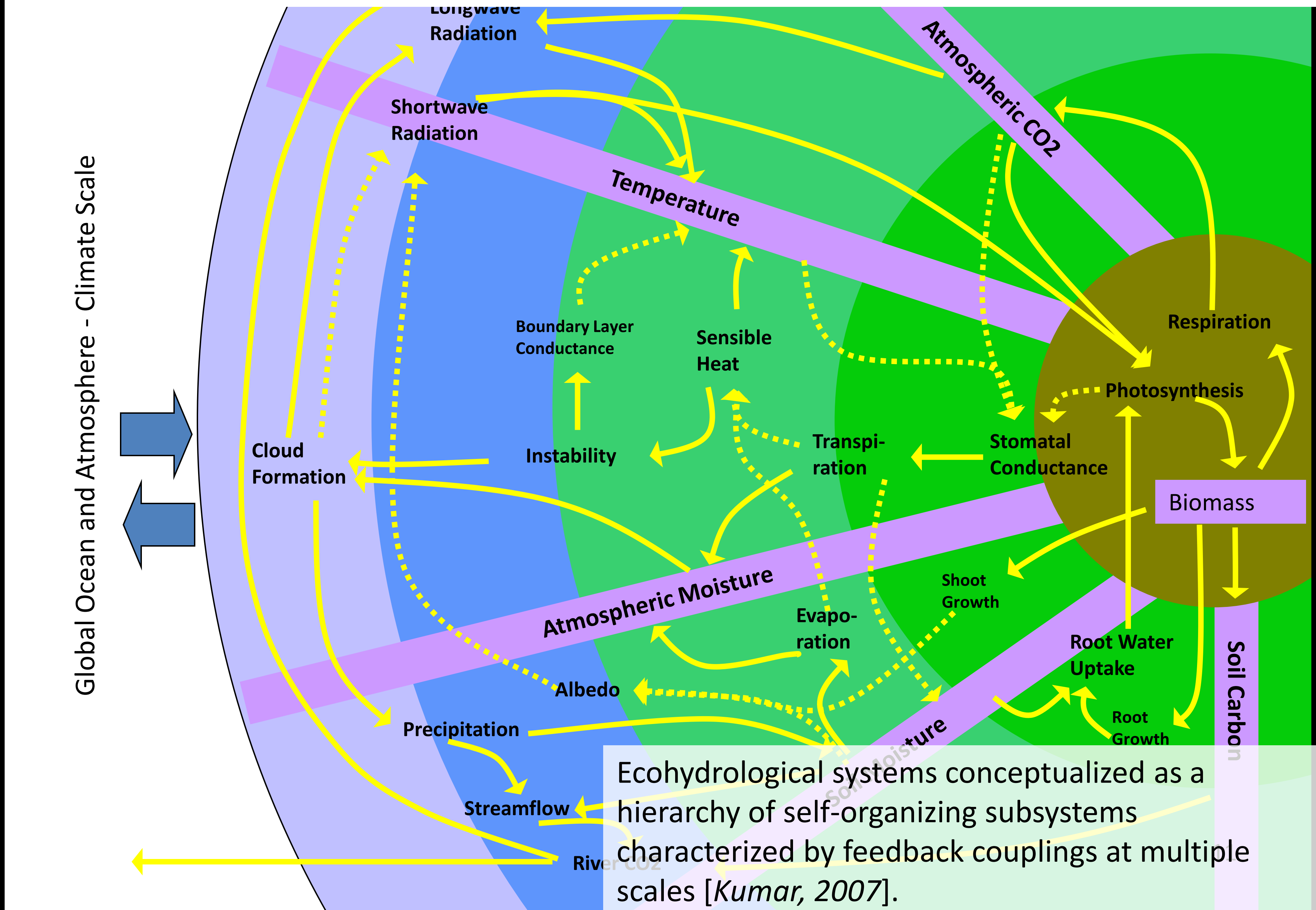
*Dynamical* DPN's vs. *Steady State* PN's

- Some connections are *Dynamical* (**dashed**) representing some kind of co-variation
- Some connections are *Steady State* (**solid**) representing an average static link
- Experiments often observe co-variation to infer process; that is a DPN.
- Models usually consider changing nodes as the 'variables' and steady nodes as 'control parameters', at a given scale and state.




Don't confuse *network dynamics with dynamical process networks*

# Example: A model linear correlation multitype DPN




Ecohydrological systems conceptualized as a hierarchy of self-organizing subsystems characterized by feedback couplings at multiple scales [*Kumar, 2007*].

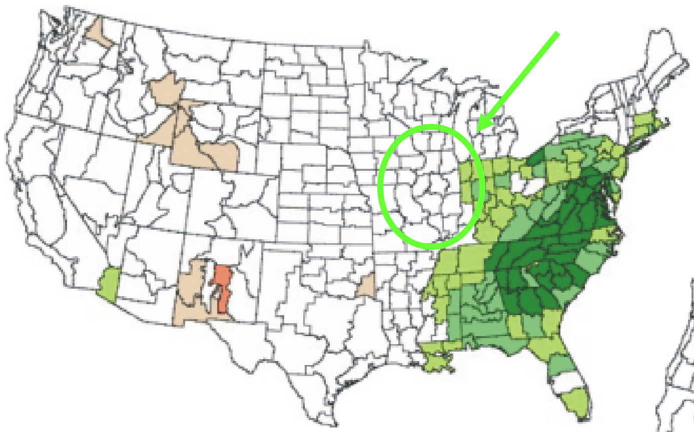
Examples where the DPN coupling thresholds distinguish system states via changing network topology

Couplings are resolved using Information Flow Statistics… ask me later!

# Hydrological drought is a state defined by Dynamical Process Network Topology decoupling

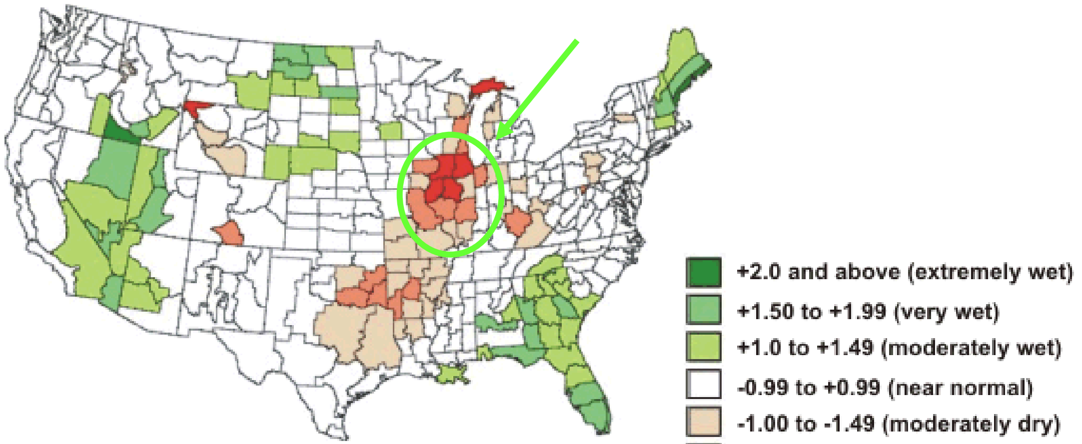
Using the Bondville FLUXNET site; a corn-soybean ecosystem

6-month SPI through the end of July 2003





6-month SPI through the end of July 2005

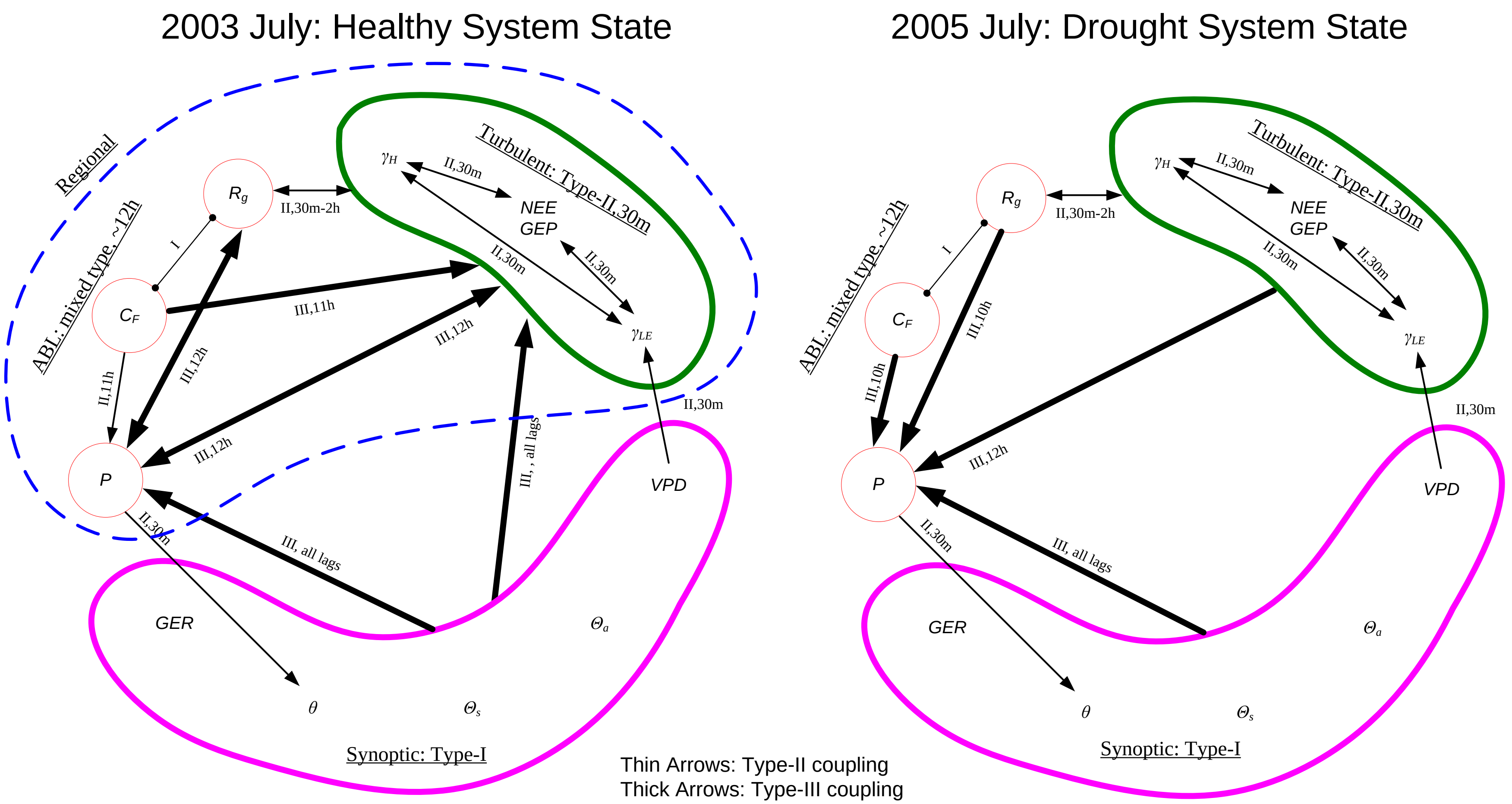





**Standardized Precipitation Index**
**http://www.drought.unl.edu/**

# Information Flow multitype DPN with network dynamics: Observed flux tower Drought state vs. Normal state

Ruddell and Kumar (2009a)




**Regional Moisture Feedback Breaks Down During Drought**

# Hypotheses and Implications

Whole-ecosystem DPN couplings derived from fine-scale dynamics, such as eddy-covariance flux observations, are a valid metric for an ecosystem's macroscale *functional* niche and role.

The existence of a significant coupling on an ecosystem's DPN during a phenostage implies sensitivity of the ecosystem to changes in the coupled subsystem that occur specifically during that phenostage, but not necessarily other phenostages.

Under climate change, ecosystems will generally transition and spatially migrate to maintain their DPN's during all phenostages.

Therefore, if climate or other forcings alter the DPN in a location, we may predict that another ecosystem that is adapted to that functional DPN role will succeed the current one.

# FLUXNET: Carbon productivity DPN coupling is sensitive to air temperature

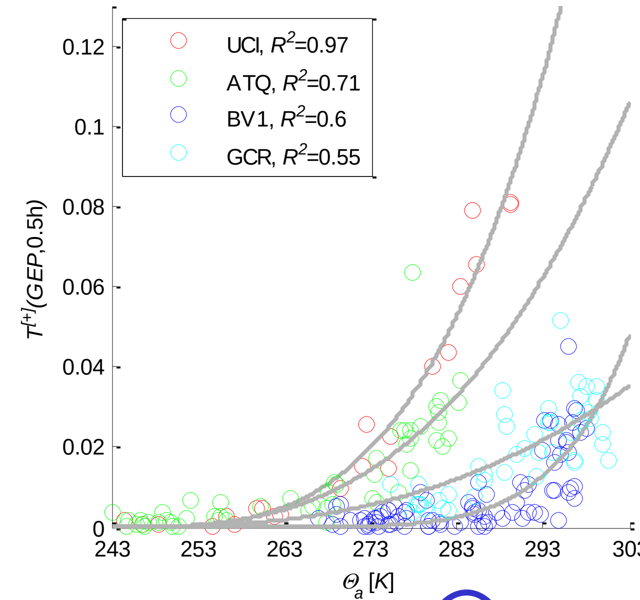
DPN coupling of ecosystem carbon productivity to other subsystems is measured by total information flow $T^{+}$. (Kumar and Ruddell 2010)

Coupling is sensitive to mean monthly air temperature.

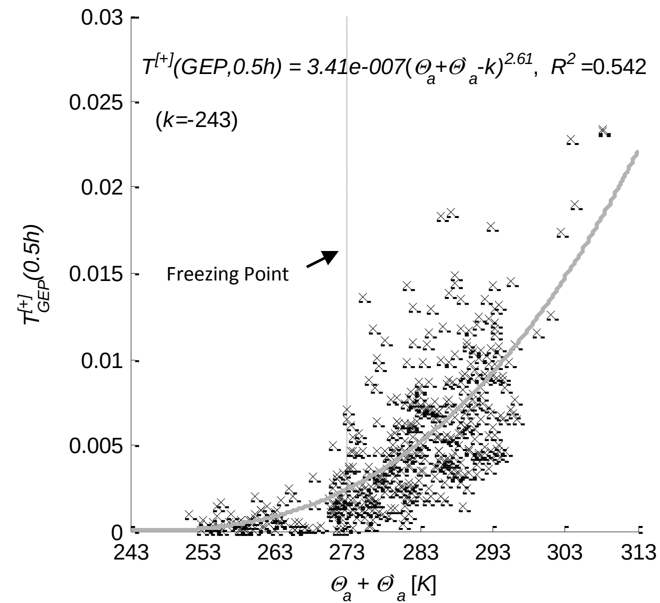
Define a new parameter, the *Thermal Offset Adaptation Temperature $\Theta_a$`*, to account for the adaptation of each ecosystem.

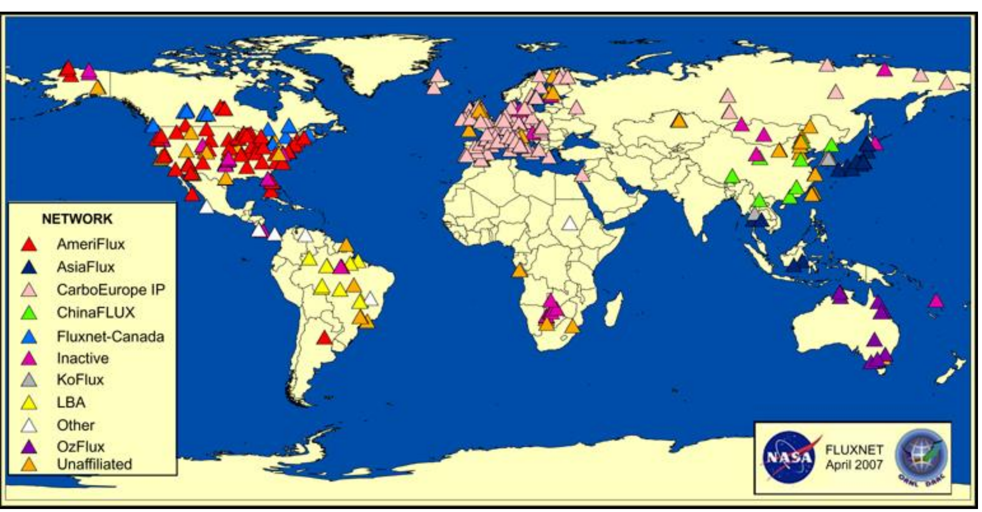
When adjusted by $\Theta_a$`, all ecosystems fall onto a single power curve relating $T^{[+]}$ to $\Theta_a$.

$$T_{GEP}^{[+]} = a \cdot (\Theta_a)^b$$




$$T_{GEP}^{[+]} = a \cdot (\Theta_a + \Theta_a` + k)^b$$

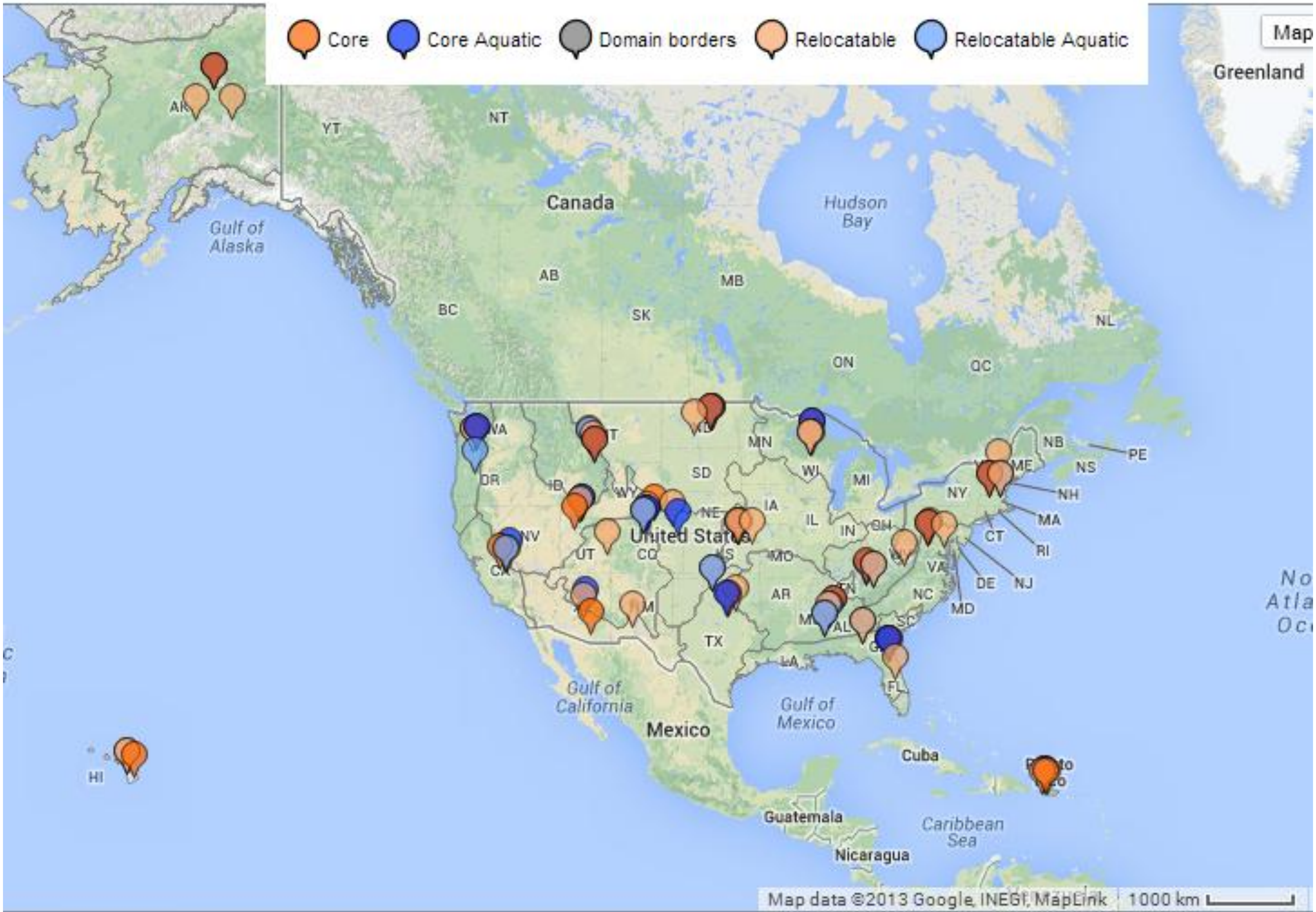



| Site | $\Theta_a$` [K] |
|---|---|
| UCI | 19 |
| ATQ | 8 |
| ARR | -8 |
| BV1 | -14 |
| GCR | -7 |
| KSC | -6 |
| TZR | -9 |
| WRC | 3 |

# NEON and Phenology Data Useful for This Analysis

# NEON Field Sites (current 2013)



neon | NATIONAL ECOLOGICAL OBSERVATORY NETWORK



Photo courtesy of Prof. Mark D. Schwartz, Dept. of Geography, UW-Milwaukee, Milwaukee, WI


USA npn
National Phenology Network
Legend
Circle size and color represent the number of sites registered with Nature's Notebook in each state. Click one to see the individual locations within a state.
nature's notebook
Canada
United States
Mexico
Hudson Bay
Manitoba
British Columbia
Saskatchewan
Ontario
Newfoundland and Labrador
Gulf of St Lawrence
New Brunswick
Prince Edward Island
Nova Scotia
Vermont
New Hampshire
Massachusetts
Rhode Island
Connecticut
New Jersey
Delaware
Maryland
District of Columbia
Gulf of Mexico
Cuba
Dominican Republic
Puerto Rico
Guatemala
Michigan
Idaho
Louisiana
Mississippi
Georgia
Arkansas
Mexico
815
443
292
290
284
285
254
222
199
198
175
171
170
152
150
148
148
116
112
112
109
106
104
91
84
79
78
76
57
55
47
44
43
40
38
32
28
28
24
18
16
14
14
13
10
7
5
1
1

Department of Organismic & Evolutionary Biology

Harvard University

# Richardson Lab

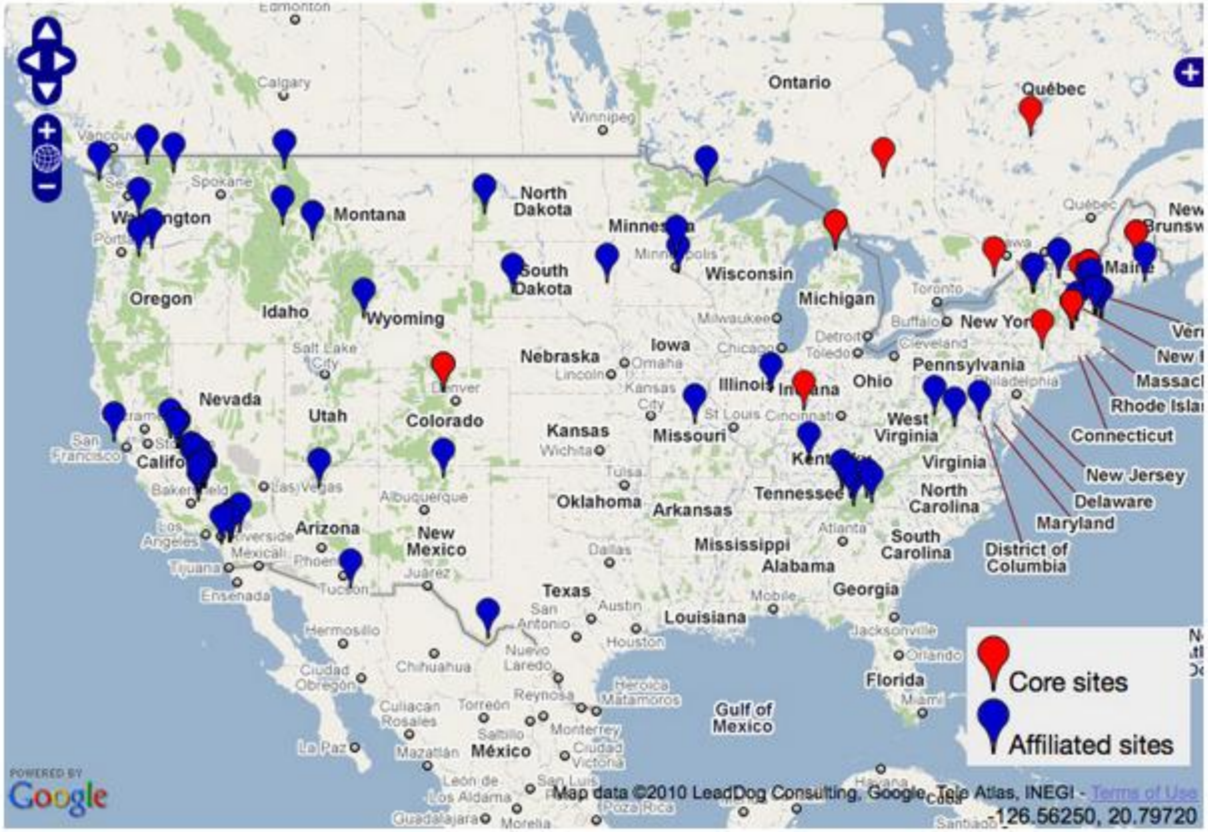
*Terrestrial Ecosystems and Global Change*



## Field Sites & Data

New England Field Sites
Gradient Studies
PhenoCam
Remote Sensing
FLUXNET
Models

## Phenocam




We initiated PhenoCam to provide automated, near-surface remote sensing of canopy phenology across the Northern Forest region of New England, upstate New York, and adjacent Canada. We began by installing high-resolution digital cameras ("webcams") at more than a dozen established research sites distributed throughout this region. With the collaboration of researchers and land

# Ongoing Work

- Actually testing these hypotheses… easier said than done!
- Using FLUXNET, mapping DPN's for all of the world's observed ecosystems.
- Developing a general principle defining phenostages using DPN functional roles rather than traditional biological metrics.

# References


Baldocchi, D. et al. (2001), FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bulletin of the American Meteorological Society*, *82*, 2415-2434.

Drewry, D. T., P. Kumar, S. Long, C. Bernacchi, X.-Z. Liang, and M. Sivapalan (2010), Ecohydrological responses of dense canopies to environmental variability: 2. Role of acclimation under elevated CO2, J. Geophys. Res., 115, G04023, doi:10.1029/2010JG001341.

Kumar, P. (2007), Variability, Feedback, and Cooperative Process Dynamics: Elements of a Unifying Hydrologic Theory, *Geography Compass*, *1/6*, 1338-1360, 10.1111/j.1749-8198.2007.00068.x.

Reichstein et al. (2005), On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biology 11, 1424-1439.

Schreiber T (2000), Measuring Information Transfer. Physical Review Letters, 85(2) 461-464.


# Related Publications


Ruddell, B.L., N. Oberg, P. Kumar, and M. Garcia (2010), Using Information-Theoretic Statistics in MATLAB to Understand How Ecosystems Affect Regional Climates, MATLAB Digest Academic Edition, February 2010, www.mathworks.com/academia.

Kumar, P. and B.L. Ruddell (2010), Information Driven Ecohydrologic Self-Organization, Entropy, 12, 2085-2096.

Ruddell, B. L., and P. Kumar (2009a), Ecohydrologic process networks: 1. Identification, Water Resour. Res., 45, W03419, doi:10.1029/2008WR007279.

Ruddell, B. L., and P. Kumar (2009b), Ecohydrologic process networks: 2. Analysis and characterization, Water Resour. Res., 45, W03420, doi:10.1029/2008WR007280.

# Abstract

It is increasingly likely that predictions of decadal climate change and land use change will yield the accurate information needed to anticipate ecosystem adaptation to human-induced change (e.g. climate variability, land use change). It is therefore essential that we develop new theories, modeling tools, and data products that are capable of predicting ecosystem adaptation to these changes, and that can anticipate how possible nonlinear thresholds will affect ecosystem structure, function, and services. This project links information about a land surface ecosystem's dynamics (e.g. eddy covariance flux tower data) from existing observational networks (e.g. FLUXNET, LTER, NEON), paired with ecosystem phenology data from the U.S. National Phenology Network (USNPN) to analyze how key dynamic couplings between ecosystems, climate, and hydrology change as ecosystems progress through successive phenological stages. The resulting dynamical process networks are quantitative graphs of the complex system's network of couplings during each phenological stage. By drawing generalizations and patterns from the study of many ecosystems, it is possible to use this theoretical framework to quantify how ecosystems are sensitive specific climate changes during specific phenological stages.

# Appendices

# Q: What is Information?

# Q: What is Information?

**A: Information is the Answer to a Question**

**Q: What is Information?**

**A: Information is the Answer to a Question**

**Q: What is our Question?**

**Q: What is Information?**

**A: Information is the Answer to a Question**

**Q: What is our Question?**

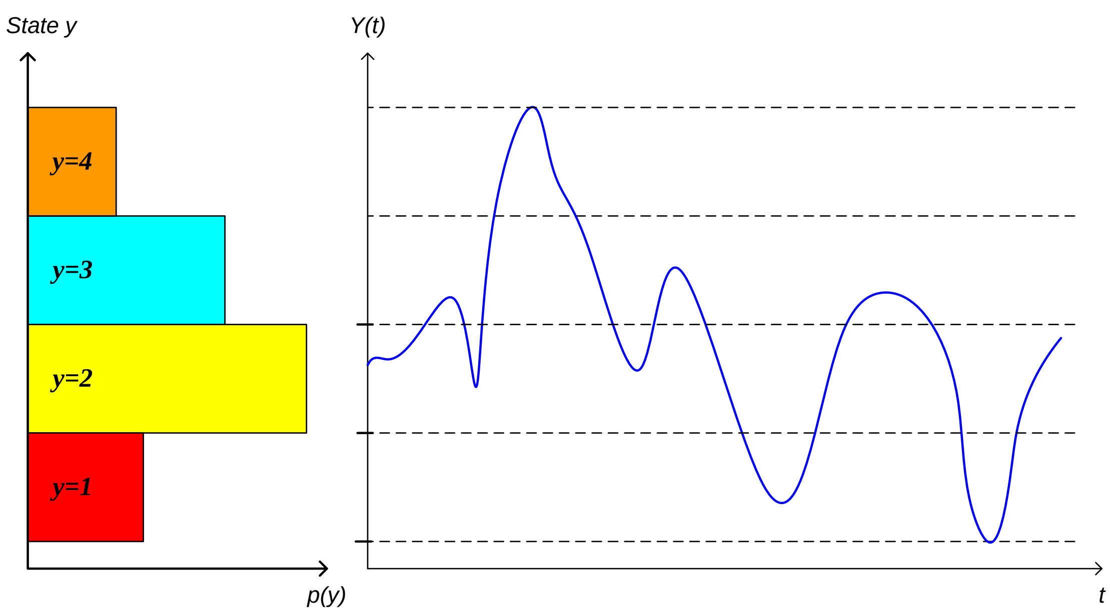
**A: "What will be the Future State of Timeseries Variable *Y(t)*?"**

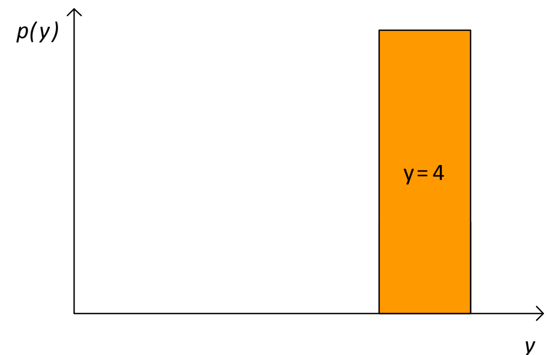
# **Shannon Entropy:** The fundamental measure of uncertainty and information

*p(y)* is the prior probability that discrete variable *Y* takes state *y.*

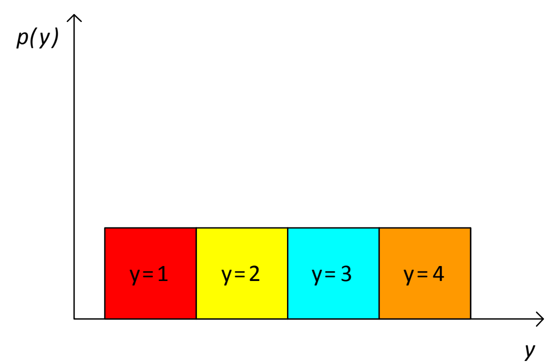
$$H(Y_t) = -\sum_{y \in Y_t} p(y) \cdot \log p(y)$$

*H($Y_t$)*, the Shannon Entropy, measures the size of the question of state; this is also the amount of information we gain when we learn the answer to the question.



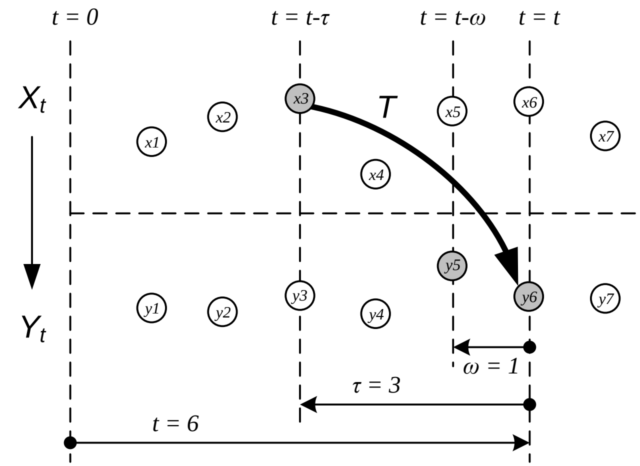

*Minimum H for 4 discrete states*
*H (Y ) = 0 bits , no uncertainty*




*Maximum H for 4 discrete states*
*H (Y ) = log $_2$(4) = 2 bits*

# How to Measure Information Flow?

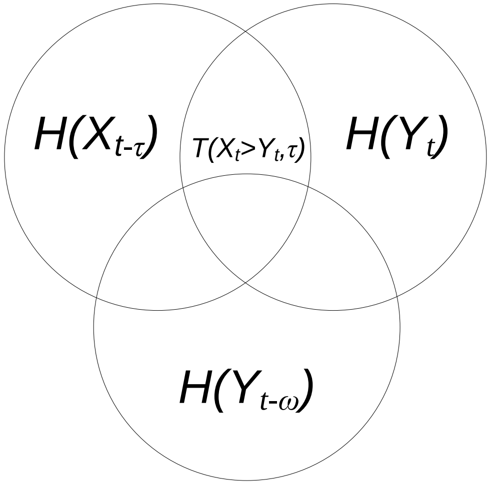
## Transfer Entropy!




- To measure directional information flow and assess timescales of flow, we need an asymmetric measure of information flow

- Thomas Schreiber [2000] introduces **Transfer Entropy** *T*, conditioning information shared by $X_t$ and $Y_t$ on $Y_t$'s history

$$T(X_t > Y_t, \tau) = \sum_{y_t, y_{t-1}, x_{t-\tau}} p(y_t, y_{t-1}, x_{t-\tau}) \log \frac{p(y_t \mid (y_{t-1}, x_{t-\tau}))}{p(y_t \mid y_{t-1})}$$



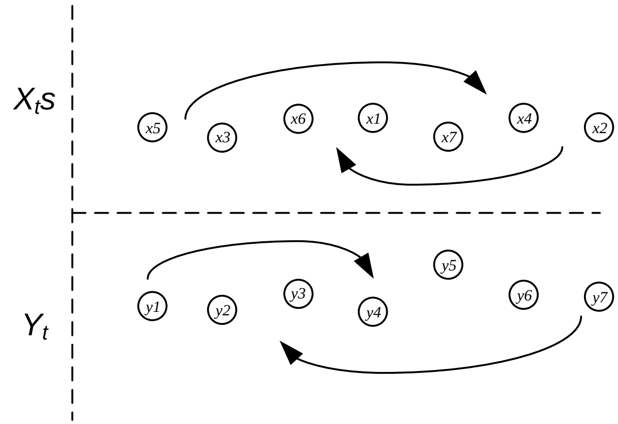

- *T* measures additional information contributed by $X_t$ across at time lag τ. Entropy reduced = information produced.

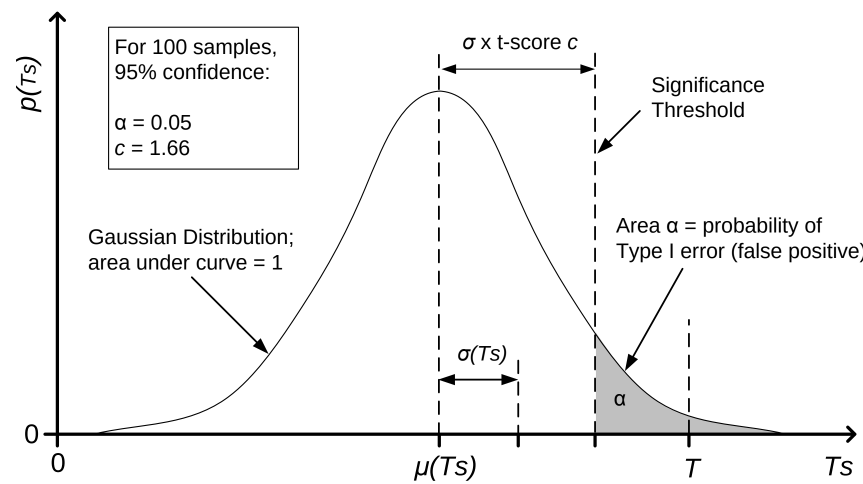
- By computing *T* across many time lags, we can assess the time scale of directional coupling from $X_t$ to $Y_t$

$$T(X_t > Y_t, \tau) = H(X_{t-\tau}, Y_{t-1}) + H(Y_t, Y_{t-1}) - H(Y_{t-1}) - H(X_{t-\tau}, Y_t, Y_{t-1})$$

# Establish Statistical Significance of Information Flow between $X_t$ and $Y_t$

- How do we decide whether *T* is large enough to represent a significant flow of information?
- Compare measured *T* against *Ts*, which is the information flow using a time-shuffled $X_t$ and $Y_t$ "bootstrapping".
- When *T* > *Ts*, a significant information flow exists; $X_t$ contributes significantly to our ability to answer questions about future states of $Y_t$.
- Robustness of results additionally ensured by quality control including testing on coupled Logistic maps, and with various *N*, *m*, and binning schemes.



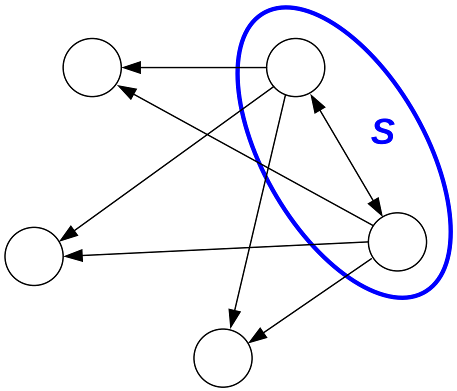
# Measures of Control and Synchronization in dynamical systems

Ecohydrologic process networks:
2. Analysis and characterization

Benjamin L. Ruddell[1] and Praveen Kumar[1]

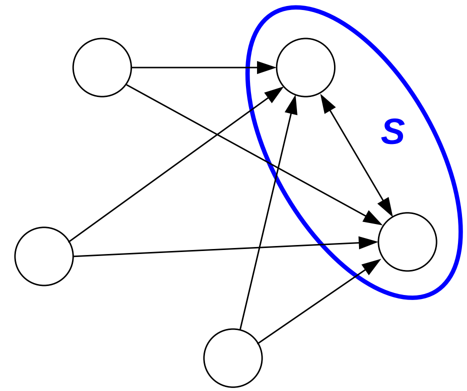
Received 11 July 2008; revised 25 November 2008; accepted

**Gross information production $T^{[+]}(S)$**
***A measure of system control exerted by S***




$$T_S^{[+]}(\tau) = \sum_{\substack{i \in S \\ z \in V}} \mathbf{A}(i, z, \tau)$$

**Gross information consumption $T^{[-]}(S)$**
***A measure of the system's control of S***




$$T^{[-]}(S, \tau) = \sum_{\substack{i \in S \\ z \in V}} \mathbf{A}(z, i, \tau)$$

**Net information production $T^{net}(S)$**

$$T^{net}(S, \tau) = T^{[+]}(S, \tau) - T^{[-]}(S, \tau)$$

**Total information production *TST(V)*** is the normalized sum of $T^{[+]}(S)$ *across all subsystems S*

**Mean System Shannon Entropy *H(V)*** is the normalized average of all subsystem Shannon Entropies *H(S)*